Reliability evaluation of incomplete AIS trajectories

Laurent Lecornu, Julien Montagner, John Puentes

Reliability Evaluation of Incomplete AIS Trajectories

L. Lecornu, J. Montagner and J. Puentes

Institut Mines Télécom; Telecom Bretagne, Département Image et Traitement de l'Information; Lab-STICC, équipe DECIDE, Brest, France
{laurent.lecornu, julien.montagner, john.puentes}@telecom-bretagne.eu

Abstract. Automatic Identification System (AIS) data are prone to alterations that impede signal storage, producing incomplete trajectories. Since trajectory analysis assumes complete AIS data sets, incomplete trajectories are discarded. Yet, those data sets contain proper values, which could be exploited. This paper describes an approach to estimate the degree of reliability assigned to missing segments of a vessel trajectory, depending on meaningful statistical variations of recorded and predicted vessel positions. The approach was tested on real AIS data of three vessels. Results suggest that reliability can be determined from vessel speed and heading variations.

Keywords: **Missing data analysis, reliability estimation, trajectory prediction, statistic distributions.**

1 Introduction

Trajectory analysis is a fundamental task to understand maritime navigation traffic, detect anomalous motion patterns, and perform surveillance according to movement prediction. For instance, trajectory anomalies were detected applying a learning framework based on reference patterns [1]. Accurate movement prediction and tracking were calculated for practical scenarios requiring shared information, reducing communication costs [2]. Also, data mining of dense trajectories sets resulted in detection of main routes, outliers, and unusual situations [3]. Data for these studies were generated by the self-reporting Automatic Identification System (AIS), progressively used in vessels. Being a non-controlled broadcasted data source, it is frequently prone to data alterations like, invalid positions, interferences, message collisions, and missing values. In this last case, while transmission or reception interruptions may occur because of technical problems or navigation in out-of-range areas, signal transmission can also be voluntarily interrupted in the vessel to stop tracking. Both situations imply that for some time AIS data are not recorded by the reception station, and the corresponding trajectory will show gaps when displayed. Since trajectory analysis algorithms assume complete AIS data sets, tacitly discarding missing data, incomplete trajectories have received little attention. Nevertheless, those incomplete data sets contain valid useful values, which could be exploited. This paper proposes an approach to estimate the degree of reliability assigned to each part of trajectories, particularly to incomplete segments, estimating if the incomplete data set can be exploited in further processing.

The developed method is based on the statistical study of each vessel trajectory, and analyses the differences between values produced by the AIS and the corresponding predicted values. Its main contribution is to assign a reliability index to trajectory segments determined by pairs of successive vessel positions, depending on statistically meaningful normal and abnormal variations of recorded and predicted data. In the rest of the paper: Section 2 presents the proposed trajectory reliability analysis approach; section 3 describes how to apply it to vessel trajectories in order to detect incomplete ones; preliminary results are discussed in section 4; section 5 outlines the contribution and perspectives of our work.

2 Analysis of missing data

Missing data in AIS-based trajectory analysis occur when no value was stored by the reception station during a given period of time, within a longer interval in which some valid data were recorded. Except for a recent study that abstracted missing data as a conditioned interpolation equivalent to a 'trip' hypothesis [4], incomplete trajectories are simply discarded before trajectory analysis. To deal with this issue we propose to analyze the reliability of each pair of successive AIS records, according to the respective update time interval defined by the AIS standard [5]. Whenever the update delay is higher than the corresponding expected interval, AIS values are missing, producing incomplete trajectory segments and raising therefore the question about how reliable those variations can be.

Two parameters are studied. First, the time elapsed between two consecutive AIS records, which is compared to the expected time between two data emissions according to the vessel speed on the first point of the considered pair of points. Second, the distance difference between the second AIS point and the predicted vessel displacement calculated using its speed and heading on the first point [2]. These two parameters permit to determine how reliable is the trajectory segment defined by a pair of AIS points. To this end, the statistical distribution of previous parameters, as a function of vessel speed and heading on the first point, were analyzed for each vessel trajectory. Considering these distributions, risk values (R_1 and R_2) are assigned to the trajectory points, defining the lack of reliability of trajectory segments as:

$$R_{1,2}\left(x\right)=\int_{x_r}^{x}-\log\left(p\left(u\right)\right)du \qquad (1)$$

where x defines the parameter value (update time interval in case 1, or prediction error in case 2) of the current point, x_r is the reference value for the given parameter according to the AIS specification with respect to the vessel speed and heading ($x_r = 0$ in the case of prediction error). Function $p(x)$ is an estimation of the corresponding parameter probability distribution, on which the occurrence rareness of any x is obtained calculating Shannon's amount of information (-log $p(x)$), permitting to detect if the current point is an outlier. Each risk factor is normalized by $T_{1,2}$, which results from the same integration but on the interval $[x_r, +\infty]$, representing the maximum risk. The global reliability index is then calculated applying a conjunctive combination of the preceding indexes, in such a manner that a weak risk index is an attenuating factor for the other risk index: $P(x) = 1 - R_1(x) \cdot R_2(x)$.

3 Experiments

The proposed method was applied to three representative incomplete cargo trajectories (Fig. 1) of the AIS_2009 data base [6]. The central figure shows three vessels trajectories identified as tr_1, tr_2, and tr_3. Color lines represent the processed reliability index (near to 0 in red, and up to 1 in green).

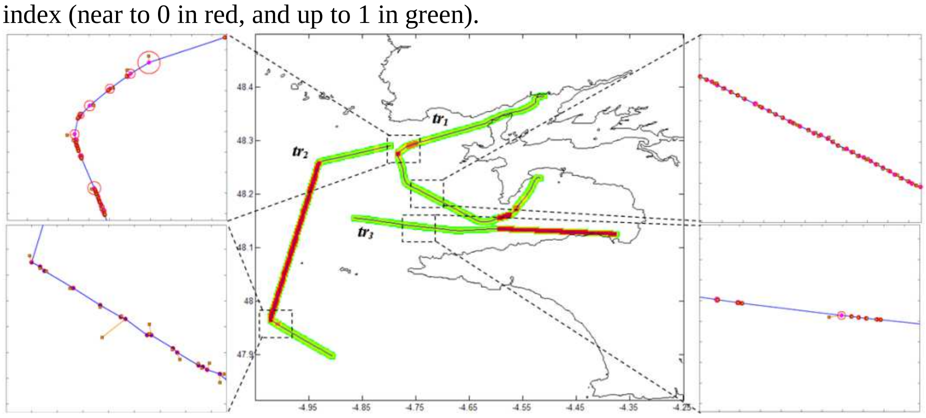

Fig. 1. Visualization of reliability evaluations for the trajectories of three vessels (tr_1, tr_2, tr_3), and examples of four different situations of data segments.

Details of reliability factors that influence the index are shown in zoom-boxes corresponding to four typical situations: top-right, short time gaps (illustrated by red circles) between acquisitions and constant vessel behavior (speed and heading); bottom-right, long time gaps, and thus larger displacement distance between acquisitions, compensated by a reliable prediction of the next position (error between predictions and corresponding points are drawn in orange) due to the straight direction of the vessel; bottom-left, quite close acquisitions, although with a high prediction error, resulting in a low reliability of the corresponding points on trajectory tr_2; top-left, both factors result in a low reliability, due to a relatively high speed of the vessel, and high prediction errors stemming from changes of heading. Both tr_2 and tr_3 present long segments where the reconstruction of the trajectory is doubtful, corresponding to time intervals where no AIS signal has been received, probably due to areas where receivers are out-of-range.

4 Discussion

Variations of AIS update time interval do not follow rigorously the AIS specifications, producing incomplete trajectory segments on which one or multiple points are missing. While some missing values approximately follow the expected displacement, others generate considerable dissimilar gaps, which are inherently related to the navigation dynamic of each vessel, during the considered trajectory. For this reason, it is proposed to base missing segments reliability estimation, on the statistical distributions of each vessel AIS data. Obtained results show a flexible reliability analysis of missing segments, detecting the lowest reliability when the vessel changes of speed or heading. It is also interesting to note that no a priori considerations about trajectory shape are required to determine the reliability index.

5 Conclusions

An approach was conceived to calculate the reliability of each point of a vessel trajectory according to the statistical significance of two parameters distributions, which depend on AIS records and associated trajectory estimations. Accordingly, each vessel data set determines individually the reliability analysis of its missing trajectory segments, instead of using only the AIS specifications and related thresholds as references. This approach can be automatically applied at the preprocessing stage to determine which incomplete trajectories could be included in a given study, depending for instance on the minimal reliability value of the trajectory. Future developments of this work include an optimization by means of distributions modeling, to identify the type of vessel based on its trajectory characteristics, and to apply the proposed approach to the complete AIS_2009 data base.

References

1. Ristic, B., La Scala, B., Morelande, M., Gordon, N.: Statistical Analysis of Motion Patterns in AIS Data: Anomaly Detection and Motion Prediction. In, Proceedings of the 11th International Conference on Information Fusion. ISIF IEEE pp. 1-7 (2008).
2. Redoutey, M., Scotti, E., Jensen, C., Ray, C., Claramunt, C.: Efficient Vessel Tracking with Accuracy Guarantees. In, Proceedings of the 8th International Symposium on Web and Wireless Geographical Information Systems. M. Bertolotto, C. Ray, X. Li (eds.), Springer-Verlag, LN series in Computer Science (LNCS 5373), Shanghai, China pp. 145-157 (2008).
3. Etienne, L., Devogele, T., Bouju, A.: Spatio-Temporal Trajectory Analysis of Mobile Objects Following the Same Itinerary. In, Proceedings of the 14th International Symposium on Spatial Data Handling, Hong Kong, pp. 86-91 (2010).
4. van Hage, W.R., Malaisé, V., de Vries, G.K.D., Schreiber, G., van Someren, M.W. Abstracting and reasoning over ship trajectories and web data with the Simple Event Model (SEM). Multimedia Tools and Applications, vol. 57, issue 1 pp. 175–197 (2012).
5. International Maritime Organization, Guidelines for the onboard operational use of shipborne Automatic Identification Systems (AIS), resolution A.917(22), 14 pages (2002).
6. http://www.chorochronos.org/?q=node/9, consulted on April 11, 2013.